



Proposal for CM Enhancements in 802.3

Manoj Wadekar
Gary McAlpine
Tanmay Gupta

Intel Corp

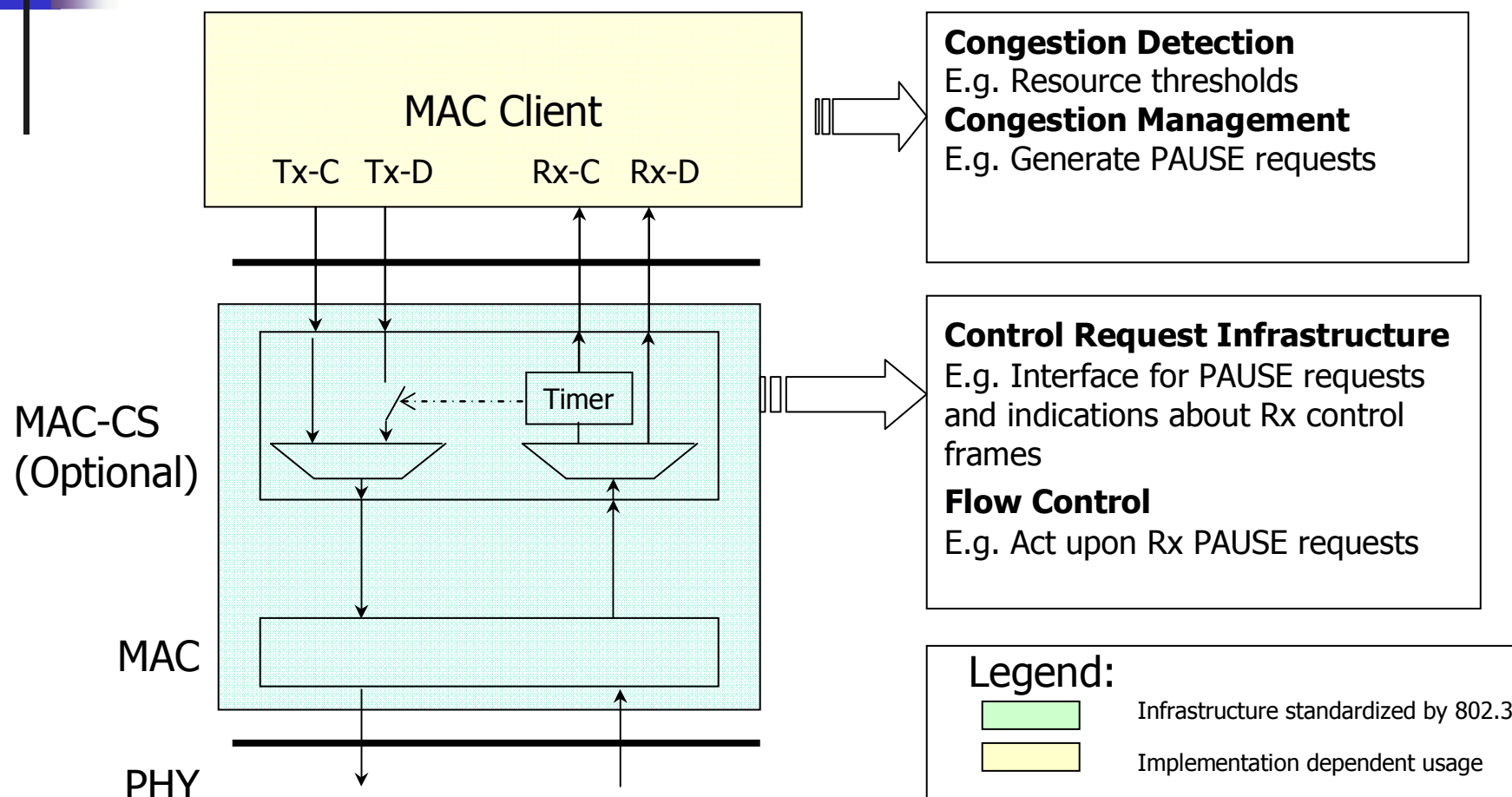
CMSG, July 2004



Agenda

- Congestion Management Model: Today
- New CM model: with Rate Control
- Summary

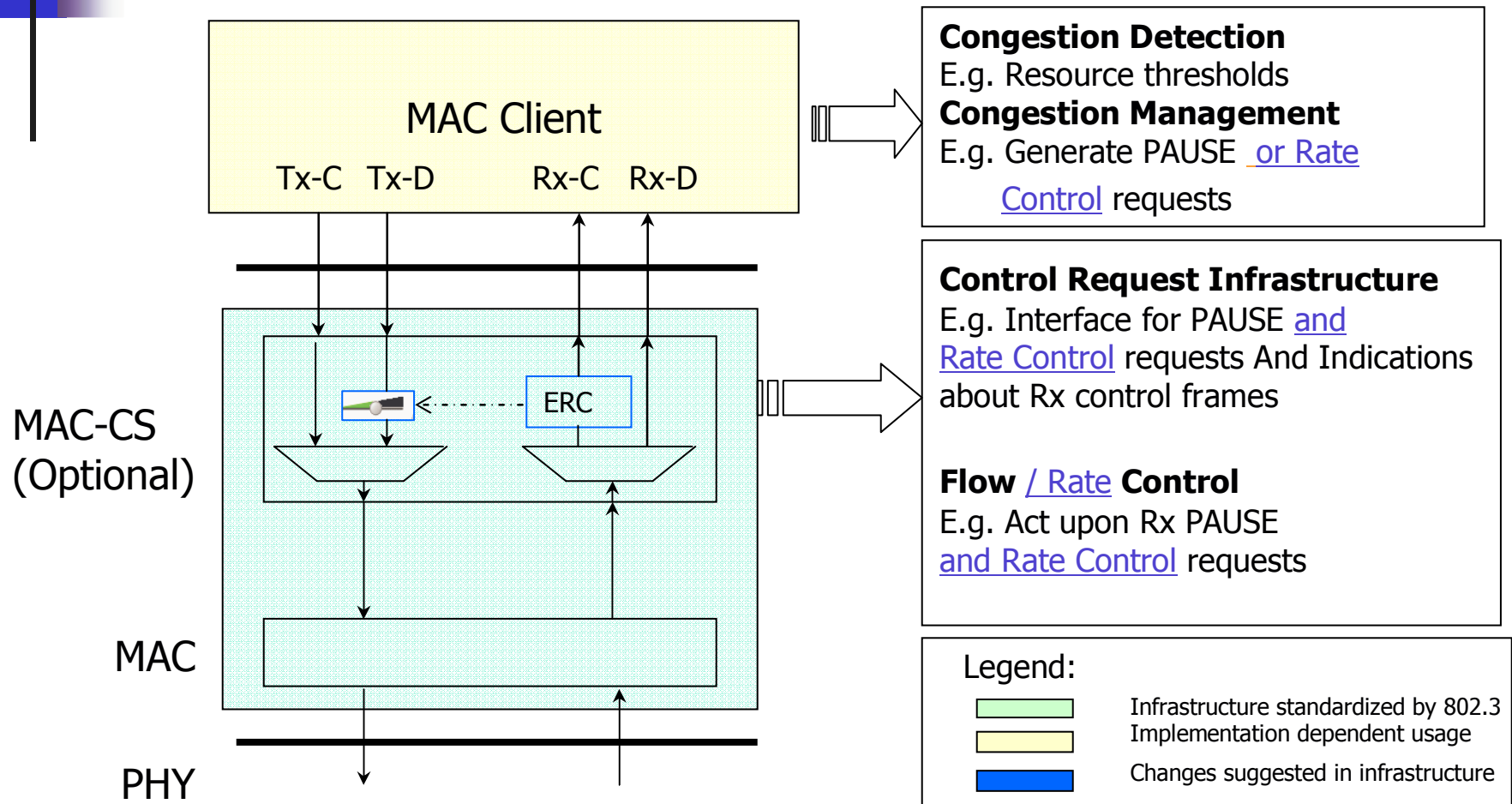
Congestion Management today



Typical response of switches today: drop the packet and let TCP adjust rate
802.3x is seldom used in typical deployments

CM: with Rate Control ..option 1

(Explicit Rate Control: ERC)



Explicit Rate Control: MAC-CS changes TX rate as per explicit request from link-peer



MAC Control Sublayer Enhancements

- MA_Control Interface
 - Request
 - Allow MAC Client to send rate control request to link partner
 - Contents of the request TBD
 - Compatible with existing MAC Control Frames
 - Should inform MAC Client about Rate change(?)
- MAC Control Sublayer
 - Define how request from link partner is handled
 - Explicit Rate Control - behavior
 - Define MAC Control Frame - enhancements
 - Granularity of the request can be discussed

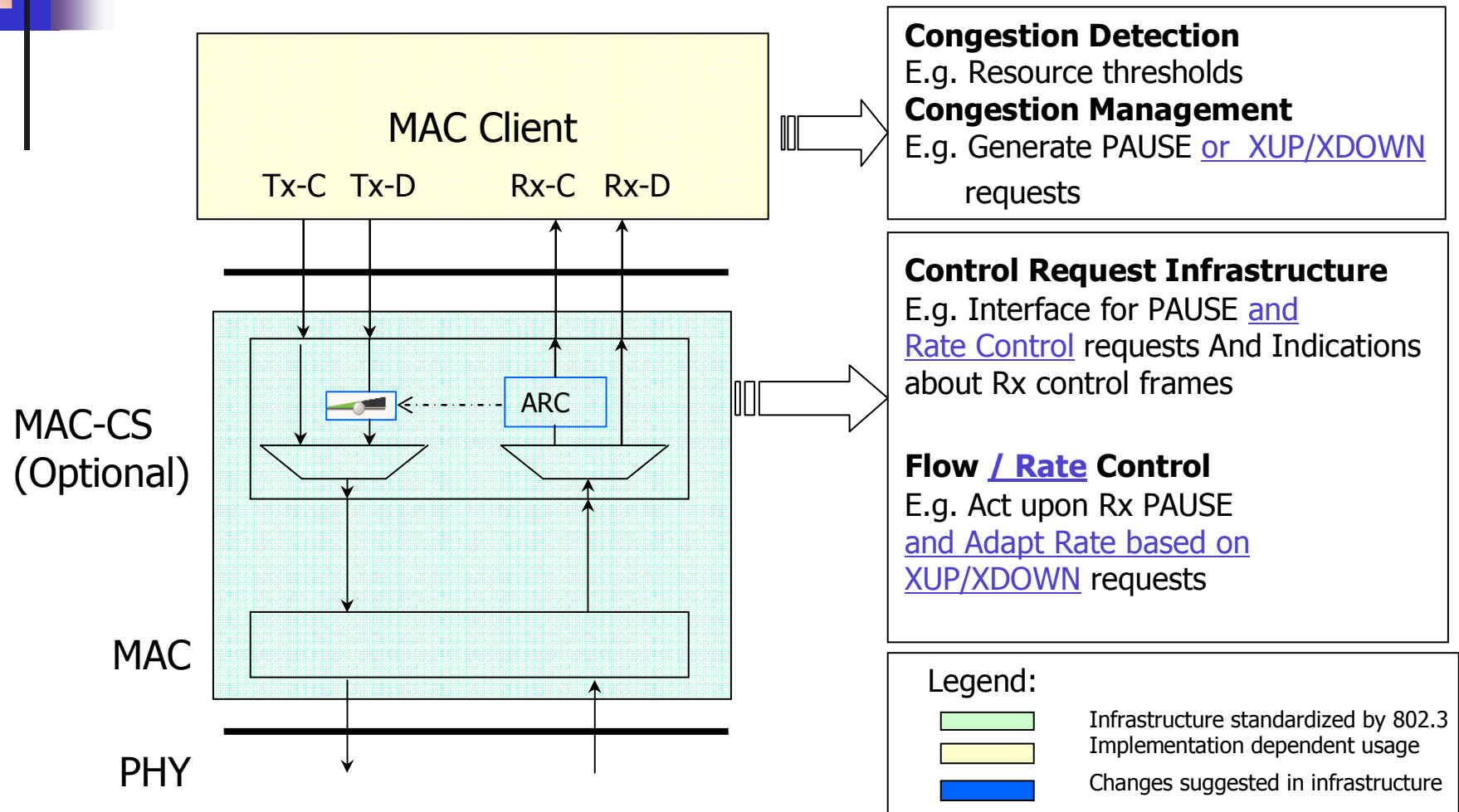


MAC Client Responsibilities

- Mechanism for detecting congestion
- Initiate CM request
 - Rate, PAUSE
- Management/Configuration
 - Enable/disable of CM functionality
 - TBD

CM: with Rate Control ..option 2

(Adaptive Rate Control: ARC)



ARC: Adaptive Rate Control: MAC-CS dynamically "adapts" rate as per XUP/XDOWN requests from link-peer



MAC Control Sublayer Enhancements

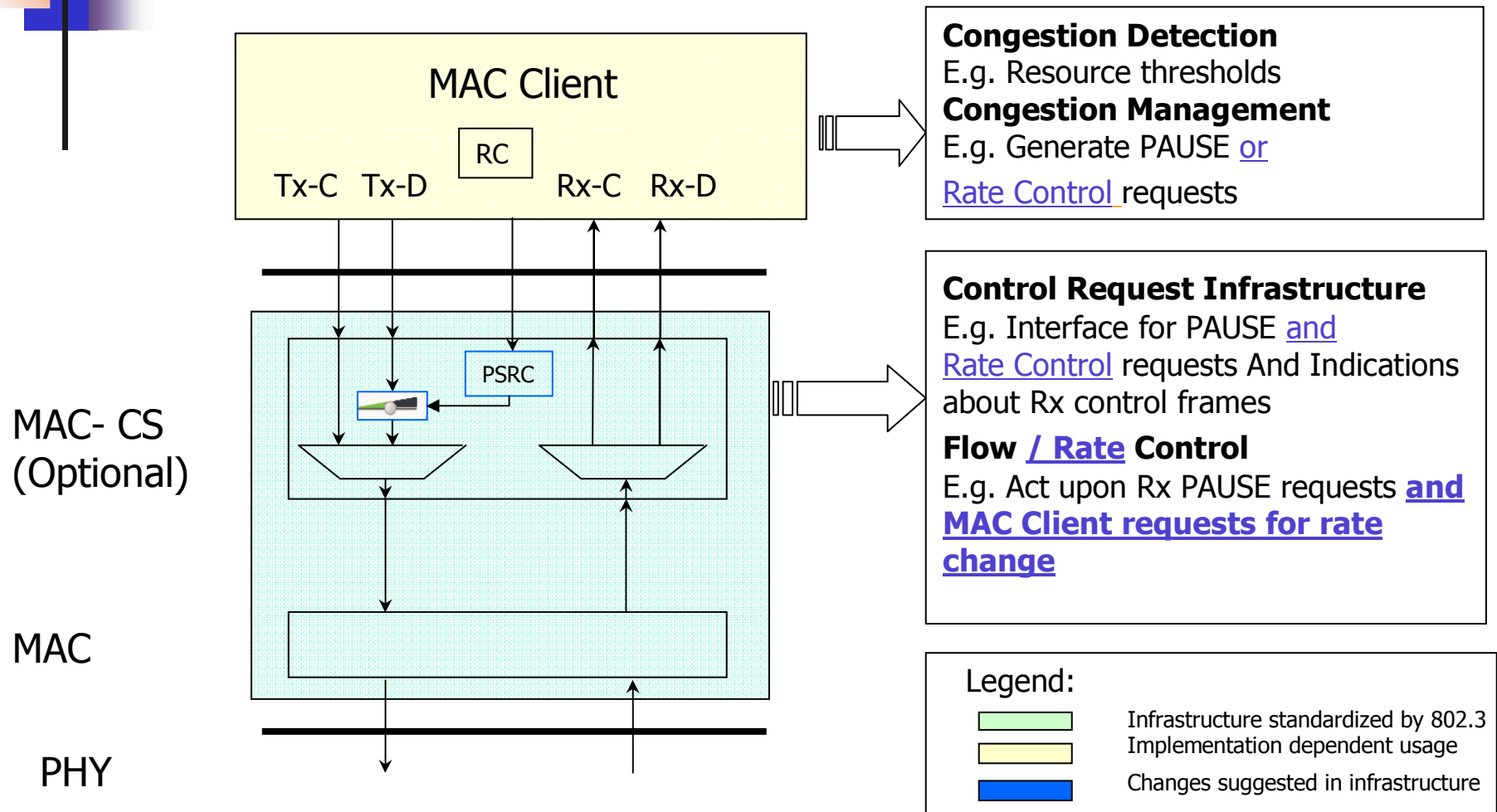
- MA_Control Interface
 - Request
 - Allow MAC Client to send request to link partner
 - Granularity of the request can be discussed
 - Compatible with existing MAC Control Frames
 - Should inform MAC Client about Rate change(?)
- MAC Control Sublayer
 - Define how request from link partner is handled
 - Adaptive Rate Control as response to XUP/XDOWN request from link partner
 - Granularity of the control TBD



MAC Client Responsibilities

- Mechanism for detecting congestion
- Initiate CM request
 - XUP/XDOWN, PAUSE
- Management/Configuration
 - Enable/disable of CM functionality
 - TBD

CM: with Rate Control ..option 3 (Pseudo-Static Rate Control: PSRC)



PSRC: Adaptation is left to MAC-client. MAC-CS provides service interface for rate change.



MAC Control Sublayer Enhancements

- MA_Control Interface

- Request

- Allow MAC Client to send request to link partner
 - Allow control of MAC-CS for changing Tx Rate
 - Contents of the request TBD
 - Compatible with existing 802.3x flow control

- Should inform MAC Client about Rate change(?)

- MAC Control Sublayer

- Define how request from MAC Client for Rate Change is handled
 - Contents of the request TBD



MAC Client Responsibilities

- Mechanism for detecting congestion
- Initiates or responds to CM request
 - CM request must be specified
- Respond to CM request with implementation specific algorithm
 - Control MAC-CS for rate change, if needed
- Management/Configuration
 - Enable/disable of CM functionality
 - TBD



Why MAC Control Messaging?

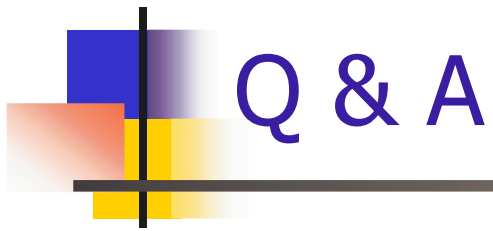
- Need CM supported on NICs & Switches
 - No current NIC standard except 802.3
- MAC Control messaging and interface:
 - Common to both NICs and Switches
 - Only small tweaks required



Summary

- Simple changes in 802.3 are possible that will enable CM capabilities in Ethernet to:
 - Reduce/avoid packet drops
 - Reduce end-to-end latency
 - Optimize throughput
- Various alternatives possible for dividing functionality between 802.3 and 802.1
 - More likely to be extensible to larger topologies
- MAC Control sub-layer CM enhancements can be compatible with 802.3x flow control

CMSG should continue to study simple rate control mechanisms for improving congestion management in Ethernet



Q & A
